COSC 223: Probability and Computing
Project 3: Randomized Data Structures and Algorithms

Due Wednesday, May 15, 11:59pm

1 Group Assignments

This is a group project: you will work with two or three other students in the class. Unlike projects 1 and 2, I have assigned groups for this project. Within each group, all group members have a similar course background (i.e., completed 311, completed 211 but not 311, currently taking 211, etc.). This is because the types of topics that are reasonable for students who have completed 311 are not the same as the topics that are appropriate for students who are currently taking 211. Your project should be at the right level for you and your group members, and should draw on knowledge from other courses that you've taken. Part of my assessment of your project will be evaluating the extent to which you've pushed yourself to make connections to what you've learned in other CS courses.

I have given each group a short list of suggested project topics that are at the right level for the members of your group. I encourage you to select a topic from this list. If all of your group members feel strongly that you want to pursue a different topic, you must let me know via email by Sunday, April 14, 5:00pm. I will not consider any off-list topics after this deadline.

Please make sure all of your group members are included on any email communications.

As with Projects 1 and 2, at the end of the project you will submit an individual reflection on your teamwork process and division of labor of your group. Building effective teamwork strategies is an important goal of this project, and your teamwork process will count towards your grade on this project.

Any proposed off-list topics: Sunday, April 14, 5:00pm
Final topic selection: Monday, April 15, 5:00pm, tell me your topic and sign up for a presentation slot on the Google sheet linked on the course web page
Reflection: Friday, May 17, 5:00pm

2 Intellectual Responsibility

You will be working with a group on this project and all aspects of the project can (and should) be done together. You may discuss the concepts involved in this project with other students in the class who are not in your group; if you do, please note at the top of your writeup with whom you consulted, and what you discussed. This project will involve finding lot of information about your topic from various online and print sources. Make sure you **cite all sources** that you used for your project. All of the writing that you produce must be your own. Learning about a data structure or algorithm from an existing source is great, but copying or paraphrasing text written by someone

else is plagiarism.

3   The Project

The goal of this project is to study in depth a randomized algorithm or data structure that's different from what we'll see in class. There are three deliverable components to your project:

- A written tutorial explaining your data structure or algorithm (Section 3.1)

- An in-class presentation at the midpoint of the project (Section 3.2)

- An implementation and experimental component and writeup (Section 3.3)

3.1   Written Tutorial

The purpose of the written tutorial is to introduce other undergraduate students to your topic. Imagine you were hired to write a textbook chapter on your data structure or algorithm. Your job in this part of the assignment is to write that chapter! Your tutorial should include (at minimum):

- A description of the problem your algorithm is trying to solve, or of the operations your data structure must provide.

- A description of how your algorithm/data structure works. You may want to provide pseudocode, show an example, or anything else that will help your readers understand your approach.

- Some analysis of the performance and correctness of your algorithm/data structure.

- A list of references of all sources you used while researching your topic and preparing your tutorial.

While you're writing the tutorial, your guiding question should be "what information would I want to know in order to understand when this algorithm is used, how it works, and how to analyze it?"

Draft due: The class meeting during which you give your presentation
Final version due: Wednesday, May 15, 11:59pm

3.2   Presentation

The last week or so of classes will be devoted to project presentations. Your presentation should be 15 minutes long (including a few minutes for questions), and we'll hear from three groups each day. The audience for your presentation is your classmates, so you should aim for your presentation to be understandable to someone who has taken 223 (and maybe 211 or 311) but is not familiar with your specific problem and algorithm/data structure.

When you give your presentation you'll be about halfway through the total time you have to work on the project, and a draft of your tutorial is due on the day you present. The presentations serve two main purposes:

- To introduce the class to a large set of problems that can be solved using randomization.

- To give you an opportunity to get feedback about your project so far and your plans for the second half of the project.

To that end, your presentation should do the following:

- Introduce the problem that you're studying

- Explain the randomized algorithm/data structure that solves this problem

- Describe the research question you're planning to study in the remainder of your project time, and what experiments you intend to run to address this question

Here are some things to think about as you're preparing your presentation; when I grade your presentation I will take into account these factors:

- Is your presentation well organized?

- Is your pacing appropriate? Did you spend an appropriate amount of time on each section: the problem, the solution, and your proposed experiments?

- Does your presentation make clear why your problem is important? Did you mention some applications that can help your audience understand why they should care about this problem?

- Are your slides clear, easy to read, not too distracting? (See my tips on designing aesthetically pleasing and helpful slides)

- Is your presentation at the right level for your audience? Does the presentation enable your audience to understand the topic?

- Have you thought through your plan for implementation and experiments? Do you have clear research questions that you're planning to study?

- Does each group member have a designated part of the presentation to deliver?

Presentation slot signup due: Monday, April 15, 5:00pm
Due: sometime during the last week(ish) of classes; groups will sign up for times

3.3   Implementation and Experiments

Your goal in this section of the project is to implement and do some sort of empirical evaluation of your algorithm/data structure. In many cases, an appropriate evaluation could involve comparing your randomized algorithm/data structure to a deterministic algorithm/data structure that solves the same problem. But feel free to be creative! We've seen in Projects 1 and 2 that sometimes experimental evaluation can tell us things that analysis can't, or can help us understand settings that are harder to analyze.

In this part of the project, you will:

- Define a clear **research question** about your algorithm/data structure.

- Implement the algorithm/data structure.

- Design and run some experiments to answer your research question.

- Write up your results and findings.

- Provide a readme file that explains how to run your code and the input, output, and function of all methods.

Research question and experimental plan due: as part of your presentation
Implementation and writeup due: Wednesday, May 15, 11:59pm

4   Timeline and Submission

There's a lot of information above, so here's a detailed timeline of when the different project components are due.

Date	Tutorial	Presentation	Experiments
Sun. 4/14, 5pm	Any proposed off-list topics (email to me)		
Mon. 4/15, 5pm		Time signup and topic selection (sign up on Google spreadsheet)	
Fri. 4/26 - Fri 5/3, in class	Draft (via submission web site*)	In-class presentation	Research question and proposed experiments
Weds. 5/15, 11:59pm	Final tutorial (via submission web site **)		Final implementation and writeup (via submission web site**)
Fri. 5/17, 5:00pm	Reflection (via Moodle)		

* Please make only ONE submission for your entire group. Individual group members should NOT make individual submissions.
** Please make only ONE submission for your entire group. Individual group members should NOT make individual submissions. Please submit all of your files (code, tutorial, writeup, readme) in a single submission. There is NO need to bundle your files into a tar, zip, rar, etc. before submitting.